Performance of Cisco IPS 4500 and 4300 Series Sensors

White Paper

September 2012

Contents

1. Introduction

1.1. Purpose

This white paper illustrates how different IPS performance metrics can be used to help customers determine the appropriate sizing for their IPS deployments. Customers can use this information to help interpret the different performance numbers that are presented in the Cisco® IPS 4500 and 4300 Series data sheets. The paper also provides detailed performance results for various testing methodologies.

1.2. Interpreting Data Sheet Metrics

The Cisco IPS 4500 and IPS 4300 Series data sheets provide several performance-related metrics, including real-world average throughput and maximum inspection throughput. These metrics and the testing methodologies behind them are detailed in this white paper.

Note: The throughput numbers in this white paper are higher than the numbers listed on the data sheets. This is because a margin is used for the data sheet numbers to ensure our appliances can achieve a performance above what is listed. The results presented in this white paper are the raw results of our testing, without the margin included in the data sheets.

The data sheets include other metrics, such as maximum connections, connections per second, and average latency. These metrics and their testing methodologies are also explained in detail in this white paper.

Table 1 includes some of the key data sheet performance metrics for Cisco IPS 4300 and IPS 4500 Series sensors.

Table 1.   Key data sheet performance metrics for Cisco IPS 4300 and IPS 4500 Series.

	IPS 4345	IPS 4360	IPS 4510	IPS 4520
Real-World Average Throughput	750 Mbps	1.25 Gbps	3 Gbps	5 Gbps
Maximum Inspection Throughput	1.8 Gbps	2.4 Gbps	5 Gbps	10 Gbps
Maximum Connections	750,000	1,700,000	3,800,000	8,400,000
Connections/Second	30,000	45,000	72,000	100,000
Average Latency (µs)	< 150	< 150	< 150	< 150

Performance is dictated by several factors, including traffic conditions on the network the IPS is deployed in, signature tuning and software signature versions. Although Cisco provides numbers based on realistic traffic mixes and network conditions, actual performance may vary.

1.3. The Importance of Accurate Performance Metrics

Network design success hinges on multiple factors, including the expected performance of the elements involved. Without proper throughput alignment, chokepoints can arise, impacting network traffic availability. Cables, interface cards, and other simple elements have fairly predictable and accurate performance guidelines. More sophisticated components, such as switches and routers, can exhibit a greater range of variance. Security elements have an equal if not greater range of performance results.

As organizations move to build higher degrees of security into their networks, predictable performance becomes even more important. Networks must be designed to assure network and application traffic will continue through traffic surges and variations - and levels typical just a few years ago are far below today's needs. The mix of traffic types has changed as well, with growth in more connection-intensive, complex applications. Given the variability of traffic mixes, the growth in throughput needs, and a cumulative gain in diverse applications plying those elements, it is increasingly critical to provide an accurate picture of how security devices will behave in these dynamic environments.

1.4. Overview of Performance Metrics

To accurately determine performance, it's vital to establish which metrics are important for the network and application environments where network security elements will reside.

Network and application success is dependent on a number of performance attributes, such as the type and location of the network deployment and the types of applications present, which are sometimes aligned with the network location.

The most frequently mentioned - and most commonly abused - performance metric is throughput. Throughput is measured in terms of traffic volume passing through a point in the network on a per-second basis. Several factors can influence the throughput of network devices, and different security vendors often define throughput differently, despite attempts to define and use common performance metrics.

Some vendors will report throughput without any inspection activity. This measurement can be used for network planning purposes - for example, if a device experiences a significant failure. Vendors generally avoid specificity in describing this value, following the assumption that it is essentially a wire equivalent through the device.

Throughput can also be described a single value based on one traffic mix or one particular protocol. For the most part, those traffic mixes, traffic change velocity, packet sizes, and protocol mixes are not described; thus, potential users have incomplete information to work with.

Intrusion prevention sensors are highly tunable and, depending on how and where they are deployed, encounter different types of traffic. Signature tuning and deployment are the most important factors that can influence the throughput of an IPS device.

Deployment-centric throughput metrics provide details of the traffic mix and relevant deployment model, allowing customers to correctly size the IPS needed for their deployment.

Latency is another performance metric describing the amount of time it takes for the traffic to pass through the device. In effect, it is a measure of the amount of time the security device must take to perform its tasks.

Connection metrics come in either as a "maximum count" or in the form of velocity representing connections per second. These metrics can significantly impact the types of applications being supported, as well as the types of devices participating.

2. IPS 4500 Product Line

The Cisco IPS 4500 Series delivers hardware-accelerated inspection, real-world performance, high port density, and energy efficiency in an expansion-ready chassis for future growth and investment protection. Its small form factor and low power consumption were specifically engineered for space-challenged data center environments.

The Cisco IPS 4500 Series provides low latency and high-availability features to meet the needs of the most demanding networks. With hardware-accelerated deep packet analysis, the Cisco IPS 4500 Series delivers multi-gigabyte performance with dedicated space available for future expansion.

The data sheet for the 4500 product line is available here:
http://www.cisco.com/en/US/products/ps12156/index.html.

3. IPS 4300 Product Line

The Cisco IPS 4300 Series scales to serve a wide range of deployment scenarios, from small offices and branch locations to enterprise data centers. The Cisco IPS 4300 Series delivers hardware-accelerated inspection performance, high port density, and energy efficiency in a 1-RU form factor.

The data sheet for the 4300 product line is available here:
http://www.cisco.com/en/US/products/ps12143/index.html.

4. Test Setup

Breaking point Storm has become the industry testing tool of choice for performance metrics. It supports a wide range of tests and is a proven solution for generating realistic traffic mixes based on hundreds of stateful applications, providing a true measurement of performance.

Figure 1 represents the logical topology used for most of the testing, including the various throughput and connection tests. The components of the test setup include a Breaking point Storm appliance, a Cisco Nexus® 7000 Series switch. and the device under testing (DUT); in this case, a 4500 or 4300 Series sensor. The Nexus 7000 Series switch provides flexibility by allowing connection of multiple DUTs to the test system with no impact on performance. The physical cabling varies depending on the DUT and the actual test being run.

The right number and type of interfaces are connected in a way that ensures that neither the DUT physical interface rate nor the switch is the bottleneck of the test.

In all the tests described below, the sensor is set up in "Inline Interface Pair" mode with one virtual sensor and default signature configuration. The default signature configuration provides basic protection against a broad range of vulnerabilities across a wide range of environments. This default signature configuration is maintained and updated by Cisco through signature updates based on new threats and changed risks associated with current threats.
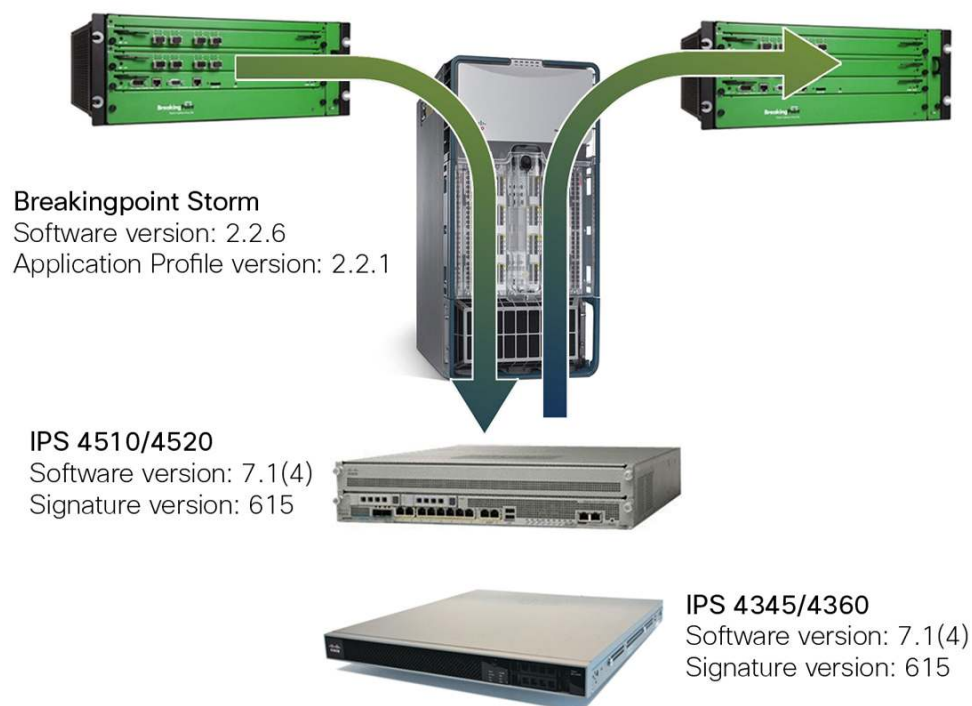
The Anomaly Detection and Global Correlation features on the IPS have been disabled, as these would not add value in a simulated performance test.

The results included in this white paper are based on the 7.1(4) code with signature release 615 for both the IPS 4510/4520 and the IPS 4345/4360.
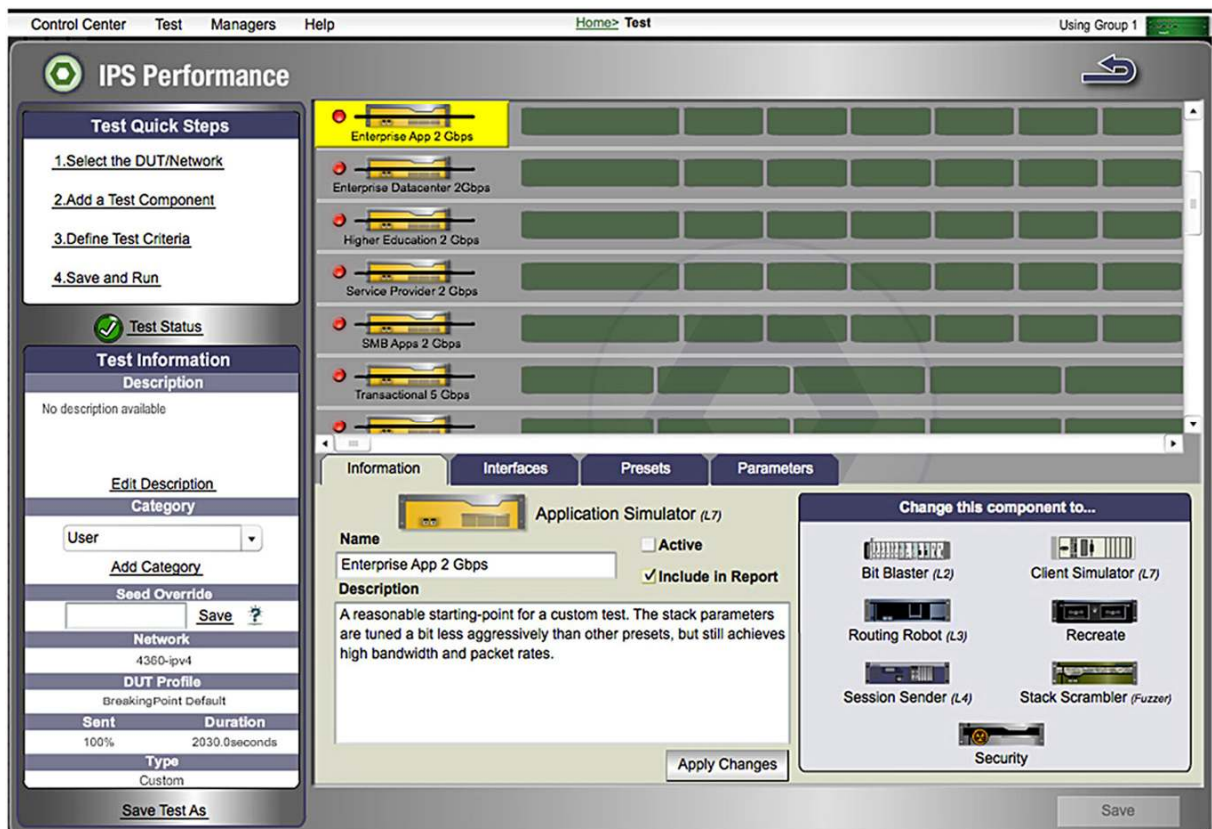
Breaking point Storm software version 2.2.6 was used for these tests. For metrics and results to be consistent across different IPS platforms, the Breaking point application profiles used for the real-world throughput tests are the profiles defined in Breaking point version 2.2.1. Breaking point regularly updates its application profiles, and Cisco regularly evaluates updating the application profiles used for real-world testing.

Results for the individual tests (especially throughput) depend on the IPS software version, signature set, and configuration/tuning. The Breaking point software version may also have an influence on the performance results.

Figure 1. Test setup

Breakingpoint Storm
Software version: 2.2.6
Application Profile version: 2.2.1

IPS 4510/4520
Software version: 7.1(4)
Signature version: 615

IPS 4345/4360
Software version: 7.1(4)
Signature version: 615

Figure 2.   Breaking point Storm throughput test configuration

5. Throughput Testing

5.1. Maximum Inspection Throughput

The purpose of the maximum inspection throughput metric is to establish and illustrate the maximum throughput the IPS can achieve while providing protection by inspecting the test traffic. This is different from a pure network throughput test, which establishes the raw packet handling capacity without inspection.

The maximum inspection throughput test Cisco uses is based on media-rich HTTP traffic with the specifications listed below.

Media-rich environments are characterized by content. Content seen on most popular websites falls on the media-rich end of the spectrum, as do video content and file transfers. If your environment is driven by access to large amounts of data and converged, immersive experiences, your environment is media-rich.

Breaking point Storm is used for this test. The application and super flow profiles have been configured as per Table 2 in order to reflect a media-rich environment. Table 3 provides the results for this test across the 4300 and 4500 platforms.

Table 2. Maximum inspection throughput characteristics

	Maximum Inspection Capacity
Traffic Protocols	HTTP
URL Length	100 bytes
Connections/Second	Low
Transactions/Connection	Two
Object Size per Transaction	Large
Average Packet Size	765 bytes
Response File Size per GET	22 KB (approximately)
Connection Duration	10 seconds (approximately)

Table 3. Maximum inspection throughput results

	IPS 4345	IPS 4360	IPS 4510	IPS 4520
HTTP - Media Rich (Mbps)	1940	2500	7530	11,140

5.2. Real-World Average Throughput

With the real-world throughput metric, Cisco not only provides a measurement of throughput using a traffic mix realistic for today's networks, but also establishes a deployment-centric testing methodology that allows customers to more accurately size IPS requirements for different deployments.

Cisco IPSs are measured using five tests that are representative of common deployment scenarios:

- Educational institution Internet edge
- Enterprise applications
- Enterprise data center
- Internet service provider (ISP) feeds
- Small and medium-sized business (SMB) or remote office flows

Each test uses an application profile (traffic mix) that is published by Breaking point.

Results of the five tests are used to generate either a range or an average, which serves as the basis for the performance ratings presented on Cisco IPS data sheets and other documentation. The intent is to give the customer a better understanding of how their unique deployment is likely to be addressed by the IPS sensor in question.

Breaking point has established these tests based on NetFlow data from networks representative of each of the five deployment scenarios.

Table 4 provides information about each of the five tests, including a graphical representation of the traffic mix. More details about the tests are available through your Cisco technical representative.

Table 5 lists the average throughput for these tests for the 4300 and 4500 product lines.

Table 4. Real-world throughput characteristics

Real-World Test Component	Protocol Mix
	P2P, HTTP, SMB/RPC, HTTPS, SMTP, CITRIX, Oracle, Other
Educational Institution Internet Edge • Based on a large university • Large amount of peer-to-peer traffic, followed by HTTP traffic	
Enterprise Applications • Wide distribution of protocols used by enterprise applications	
Enterprise Data Center • Traffic is mostly either file transfer, database transactions, or HTTP	
ISP Feeds • Collected from a well-known service provider • Mostly HTTP text/data, with a fair amount of audio and video transfers, followed by peer-to-peer traffic	
SMB or Remote-Office Flows • Mostly HTTP, with voice, database transactions, and file transfers	

Table 5. Real-world average throughput results

	IPS 4345	IPS 4360	IPS 4510	IPS 4520
Real-World Average (Mbps)	1060	1320	3340	5340

6. Latency Testing

6.1. Overview

Delays in traffic can trigger timeout conditions, which in turn can cause critical applications to fail. In some cases, time-to-live values may trigger traffic to be re-sent, potentially exacerbating challenging traffic problems. For these reasons, latency is an important consideration for an inline network security device.

Most vendors also hide latency measures. Some will make no reference to them at all, or may simply deliver a value without any kind of reference as to how it was arrived at. One vendor simply references 1518-byte packets. Others say nothing.

Generally, unless a device is deployed at a slow Internet edge, anything approaching a full millisecond is undesirable. However, even those values may be meaningless if the conditions in which they are determined are not similar to the environment in which the device will be deployed.

6.2. Latency Distribution

Cisco uses RFC 2544 (Benchmarking Methodology for Network Interconnect Devices) to measure the latency through IPS appliances. RFC 2544 defines a number of tests that may be used to describe the performance characteristics of a network interconnecting device.

The Breaking point Storm appliance has a built-in RFC 2544 test, which can be used to establish the minimum, average, and maximum latency for a given frame size and frame rate. The test also identifies the fastest rate at which frames of a specific size were transmitted. RFC 2544 specifies frame sizes of 64, 128, 256, 512, 1024, and 1500 bytes. While the latency will typically be lower for smaller packets (at a lower rate), the most relevant frame sizes are 256, 512, and 1024 bytes. The average latency at a rate close to the upper boundary for the DUT at the given frame size is provided in Tables 6-9.

On the Cisco IPS 4345 and 4360, two Gigabit Ethernet interfaces are used to establish the latency. On the Cisco 4510 and 4520, two 10 Gigabit Ethernet interfaces are used.

The data sheet latency numbers (< 150 µs for IPS 4345/4360 and IPS 4510/4520) are based on the average of the listed averages for different frame sizes.

This test is solely intended to establish the latency of a device. The given frame rate does not represent the maximum throughput.

Table 6. Average latency for IPS 4345

IPS 4345		
Frame Size (bytes)	**DUT Frame Rate (Mbps)**	**Average Latency (µs)**
256	650	148.48
512	700	143.00
1024	700	156.28

Table 7. Average latency for IPS 4360

IPS 4360		
Frame Size (bytes)	**DUT Frame Rate (Mbps)**	**Average Latency (µs)**
256	725	103.67
512	875	104.20
1024	975	130.51

Table 8. Average latency for IPS 4510

IPS 4510		
Frame Size (bytes)	**DUT Frame Rate (Mbps)**	**Average Latency (µs)**
256	800	45.21
512	1600	74.91
1024	3000	55.03

Table 9. Average latency for IPS 4520

IPS 4520		
Frame Size (bytes)	**Frame Rate (Mbps)**	**Average Latency (µs)**
256	800	39.00
512	1600	42.55
1024	3200	54.74

7. Connection Testing

7.1. Overview

The number of connections between hosts the system can track, and the rate at which those connections are established, can be critical for certain application types - and even for the types of devices involved. While sessions are generally the same as TCP connections, they have different meanings for different application types and uses. This is why most network equipment focuses on connections instead of sessions.

Connection-related metrics are most critical in transactional environments where many connections are established from a large number of hosts. Database transactions are a common example of this.

IPS vendors that understand network traffic and its relationship with connections will publish and explain these values. Unfortunately, this is rare.

7.2. Maximum Concurrent Connections

The maximum number of concurrent connections is an especially important metric for data center deployments and other application-heavy environments where the IPS must keep track of a very high number of transactions. The longer connections are kept active by the applications on the network, the higher the importance of this metric.

Table 10.   Maximum concurrent connections

	IPS 4345	IPS 4360	IPS 4510	IPS 4520
Maximum Concurrent Connections	750,000	1,730,000	3,870,000	8,600,000

7.3. Maximum Connections per Second

A maximum number holds some value, but is incomplete when building a network. Throughput and connections rarely occur in a slow, steady climb. For these reasons, the number of connections per second is important. If "bursty" traffic connections occur naturally in the deployment environment, the maximum value may not really matter, as the velocity of the growing connection rates will cause problems far before the theoretical maximum is ever reached.

Table 11.   Maximum connections per second

	IPS 4345	IPS 4360	IPS 4510	IPS 4520
Maximum Connections per Second	30,000	45,000	75,000	100,000

8. Conclusion

IPS performance metrics consist of several components, such as throughput, latency, and connection-related metrics.

Most IPS vendors do not report all IPS performance metrics, or the methodology used to get these performance metrics. Throughput numbers are often limited to either a pure network throughput number without any inspections or a single traffic standard reported without any explanation of their methodology.

Cisco publishes two throughput metrics for the IPS 4300 and 4500 Series: maximum inspection throughput and real-world average throughput. While the first metric establishes the maximum throughput that can be achieved based completely on the inspection of HTTP traffic, the second metric is based on a broad mix of traffic, with components representing different traffic mixes and deployments.

Cisco also describes our various testing methodologies, allowing customers to better choose the right IPS for their network environment and traffic mix. While some vendors treat customer deployments as if they were liabilities "outside the control" of the vendor, we see this as part of our partnership with our customers. Each individual real-world component throughput result is available to your Cisco technical representative for proper sizing needs. Cisco also publishes detailed results and methodology for latency and connection tests.

The Cisco IPS 4510 delivers a maximum inspection throughput of over 5 Gbps and a real-world average inspection throughput of 3 Gbps. The Cisco IPS 4520 delivers a maximum inspection throughput of over 10 Gbps and a corresponding real-world average of 5 Gbps.

CISCO

Printed in USA

C11-716084-00   09/12